

MAY 23 2005

PTO/SB/08A (10-01)

Approved for use through 10/31/2002. OMB 0851-0031

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449A/PTO

# INFORMATION DISCLOSURE STATEMENT BY APPLICANT

Sheet

1

of

1

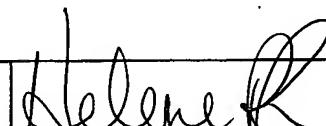
**Complete if Known**

Application No.	10/731,916
Filing Date	December 10, 2003
First Named Inventor	Mark Pearson
Art Unit	2183
Examiner Name	Not Yet Known
Attorney Docket Number	24207-10064

U.S. PATENT DOCUMENTS				
Examiner Initials*	Cite No. <sup>1</sup>	Document No.	Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document
HRR	A1	US-2001/056418 A1	12-27-2001	Youn, Seok Ho
HRR	A2	US-6,271,840 B1	08-07-2001	Finseth et al.

FOREIGN PATENT DOCUMENTS				
Examiner Initials*	Cite No. <sup>1</sup>	Foreign Patent Document	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document
HRR	B1	EP 0964341 A2	12-15-1999	Nippon Telegraph and Telephone Corporation
HRR	B2	WO 01/13273 A2	02-22-2001	North Carolina State University

OTHER REFERENCES – NON-PATENT LITERATURE DOCUMENTS				
Examiner Initials*	Cite No. <sup>1</sup>	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published		
HRR	C1	International Search Report and Written Opinion, PCT/US2004/038559, March 16, 2005.		

Examiner Signature		Date Considered	5/30/06
--------------------	---	-----------------	---------

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609.

Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup>Applicant's unique citation designation number (optional). <sup>2</sup>See Kinds Codes of USPTO Patent Documents at [www.uspto.gov](http://www.uspto.gov) or MPEP 901.04. <sup>3</sup>Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). <sup>4</sup>For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. <sup>5</sup>Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST. 16 if possible. <sup>6</sup>Applicant is to place a check mark here if English language Translation is attached.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form PTO/SB/08b

## INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet

1

of

1

**Complete if Known**

<i>Application Number</i>	10/731,916
<i>Filing Date</i>	December 10, 2003
<i>First Named Inventor</i>	Pearson et al
<i>Art Unit</i>	2183
<i>Examiner Name</i>	Unassigned

Attorney Docket Number 53051/293400

**NON PATENT LITERATURE DOCUMENTS**

Examiner Initials *	Cite No. <sup>1</sup>	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T <sup>2</sup>
HRR	1	KUSHMERICK, NICHOLAS, "Wrapper induction: Efficiency and expressiveness," Artificial Intelligence, 2000, pp. 15-88, 118, Elsevier Science B.V.	
HRR	2	BRIN, SERGEY et al., "The Anatomy of a Large-Scale Hypertextual Web Search Engine," 1998, Computer Science Department, Stanford University, Stanford, CA.	
HRR	3	LAENDER, ALBERTO et al., "A Brief Survey of Web Data Extraction Tools," 2002, Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte MG Brazil.	
HRR	4	KUSHMERICK, NICHOLAS, "Finite-state approaches to Web information extraction," 2002, Computer Science Department, University College Dublin.	
HRR	5	KUSHMERICK, NICHOLAS et al., "Adaptive Information extraction: Core technologies for information agents," 2002, Computer Science Department, University College Dublin.	
HRR	6	CHANG, CHIA-HU et al., "IEPAD: Information Extraction Based on Pattern Discovery," 2001, Dept. of Computer Science and Information Engineering, National Central University, Chung-Li, Taiwan.	
HRR	7	MUSLEA, ION et al., "Hierarchical Wrapper Induction for Semistructured Information Sources," 1999, pp. 1-27, Kluwer Academic Publishers, the Netherlands.	
HRR	8	HSU, CHUN-NAN et al., "Generating Finite-State Transducers for Semi-Structured Data Extraction from the Web," Information Systems, 1998, pp. 521-538, Vol. 23, No. 8, Elsevier Science Ltd., Great Britain.	
HRR	9	CRESCENZI, VALTER et al., "RoadRunner: Towards Automatic Data Extraction from Large Web Sites," Proceedings of the 27 <sup>th</sup> VLDB Conference, 2001, Rome, Italy.	
HRR	10	FREITAG, DAYNE et al., "Boosted Wrapper Induction," 2000, American Association for Artificial Intelligence.	
HRR	11	MICHAEL WHITE et al., "Multidocument Summarization via Information Extraction," First International Conference on Human Language Technology Research (HLT), 2001.	
HRL	12	J.-M. DELORT et al., "Enhanced Web Document Summarization Using Hyperlinks," HT'03, August 26-30, 2003, Nottingham, United Kingdom	

Examiner Signature

*Selene RR*

Date Considered

3/30/04

\*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

<sup>1</sup>Applicant's unique citation designation number (optional). <sup>2</sup>Applicant is to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 120 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.